



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 1999

The yeast expression system for recombinant glycosyltransferases

Malissard, Martine ; Zeng, Steffen ; Berger, Eric G

Abstract: Glycosyltransferases are increasingly being used for in vitro synthesis of oligosaccharides. Since these enzymes are difficult to purify from natural sources, expression systems for soluble forms of the recombinant enzymes have been developed. This review focuses on the current state of development of yeast expression systems. Two yeast species have mainly been used, i.e. *Saccharomyces cerevisiae* and *Pichia pastoris*. Safety and ease of fermentation are well recognized for *S. cerevisiae* as a biotechnological expression system; however, even soluble forms of recombinant glycosyltransferases are not secreted. In some cases, hyperglycosylation may occur, *P. pastoris*, by contrast, secrete soluble orthoglycosylated forms to the supernatant where they can be recovered in a highly purified form. The review also covers some basic features of yeast fermentation and describes in some detail those glycosyltransferases that have successfully been expressed in yeasts. These include 1,4galactosyltransferase, 2,6sialyltransferase, 2,3sialyltransferase, 1,3fucosyltransferase III and VI and 1,2mannosyltransferase. Current efforts in introducing glycosylation systems of higher eukaryotes into yeasts are briefly addressed

DOI: <https://doi.org/10.1023/a:1007055525789>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-156317>

Journal Article

Published Version

Originally published at:

Malissard, Martine; Zeng, Steffen; Berger, Eric G (1999). The yeast expression system for recombinant glycosyltransferases. *Glycoconjugate journal*, 16(2):125-139.

DOI: <https://doi.org/10.1023/a:1007055525789>



The yeast expression system for recombinant glycosyltransferases

Martine Malissard, Steffen Zeng, Eric G. Berger*

Institute of Physiology, University Zurich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

Glycosyltransferases are increasingly being used for *in vitro* synthesis of oligosaccharides. Since these enzymes are difficult to purify from natural sources, expression systems for soluble forms of the recombinant enzymes have been developed. This review focuses on the current state of development of yeast expression systems. Two yeast species have mainly been used, i.e. *Saccharomyces cerevisiae* and *Pichia pastoris*. Safety and ease of fermentation are well recognized for *S. cerevisiae* as a biotechnological expression system; however, even soluble forms of recombinant glycosyltransferases are not secreted. In some cases, hyperglycosylation may occur. *P. pastoris*, by contrast, secrete soluble orthoglycosylated forms to the supernatant where they can be recovered in a highly purified form.

The review also covers some basic features of yeast fermentation and describes in some detail those glycosyltransferases that have successfully been expressed in yeasts. These include β 1,4galactosyltransferase, α 2,6sialyltransferase, α 2,3sialyltransferase, α 1,3fucosyltransferase III and VI and α 1,2mannosyltransferase. Current efforts in introducing glycosylation systems of higher eukaryotes into yeasts are briefly addressed.

Keywords: *Saccharomyces cerevisiae*, *Pichia pastoris*, β 1,4galactosyltransferase, α 2,6sialyltransferase, α 2,3sialyltransferase, α 1,3fucosyltransferase, α 1,2mannosyltransferase, glycosylation engineering

Introduction

Heteroglycans are macromolecules which exert multiple biospecific functions by virtue of their highly ordered structures [1,2]. In recent years functions associated with their expression on the cell surface have attracted much attention because of their important involvement in cell adhesion and recognition. The paradigms for such interactions are the selectin-mediated processes [3], host-pathogen recognition for cell-cell adhesion [4], lectin-mediated endocytosis [5] and carbohydrate-mediated immune recognition [6] for glycoprotein binding.

In some instances these interactions have been proposed to be useful therapeutic targets: Inhibition of binding could prevent invasion by pathogens [7], alleviate the acute xenograft rejection [8], or the inflammatory response [9]. First generation inhibitors in general are molecules comprising the binding domain of the ligand. In the case of oligosaccharide/protein interaction the choice may be either the peptide moiety involved in binding of the oligosaccharide [10] or the oligosaccharide itself. The difficulties in obtaining sufficient quantities (i.e., gram scale) of the desired glyco-compound are notorious. Although enormous

progress has been made in the chemical synthesis of heterooligosaccharides [11], their synthesis at industrial scale still appears beyond any reasonable cost-benefit ratio. As an alternative to chemical synthesis chemo-enzymic or even enzymic synthesis is now within the grasp of possibilities.

The molecular basis for these highly specific interactions is the structural complexity of the oligosaccharide ligand. *In vivo* these are assembled by a specific class of enzymes, the glycosyltransferases. While these enzymes have been studied biochemically in some detail a quarter of a century ago (for reviews, see [12]), molecular cloning and heterologous expression has significantly advanced our knowledge (for reviews, see [13,14]).

This advance concerns mainly two topics: First, most of these enzymes belong to gene families of enzymes whose activities are usually related by their specificity for the donor substrate and a certain degree of structural homology (Table 1). Second, heterologous expression not only permits determination of their fine structural specificity towards the acceptor substrate (which may be of critical importance for subtle biosynthetic differences *in vivo*) but allows an override of their restricted specificity by using a large excess of recombinant enzyme. Thus recombinant glycosyltransferases can be used for the synthesis of non-natural glycosides [29,30].

*To whom correspondence should be addressed. Tel: 41-1-63-55070; Fax: 41-1-635-6814; Email: egberger@physiol.unizh.ch

Table 1. Gene families of mammalian glycosyltransferases

<i>Donor substrate</i>	<i>Trivial name</i>	<i>Abbreviation</i>	<i>Linkage type formed</i>	<i>Acceptor substrates</i>	<i>References</i>
CMPNeuAc	Sialyl-transferase	STX ² Acc ³ n ⁴	$\alpha 2 \rightarrow 3$ $\alpha 2 \rightarrow 6$ $\alpha 2 \rightarrow 8$	β Gal α NeuAc β GlcNAc α GalNAc	[15,16]
UDPGal	Galactosyl-transferase	$\beta 1,4$ Gal-T n ⁴	$\beta 1 \rightarrow 4$ Type 2 linkage ¹	β GlcNAc β Glc	[17,18]
UDPGal	Galactosyl-transferase	$\beta 1,3$ Gal-T	$\beta 1 \rightarrow 3$ Type 1 linkage ¹	β GlcNAc α GalNAc	[19,20]
GDPFuc	Fucosyl-transferases	$\alpha 1,2$ Fuc-T	$\alpha 1 \rightarrow 2$	β Gal	[21]
GDPFuc	Fucosyl-transferases	$\alpha 1,3$ Fuc-T n ⁴	$\alpha 1 \rightarrow 3$	β GlcNAc	[22]
UDPGlcNAc	N-acetyl-glucosaminyl-transferases	GlcNAc-T n ⁴	$\beta 1 \rightarrow 2,3,4,6$	α Man	[23]
UDPGalNAc	N-acetyl-galactosaminyl-transferase	GalNAc-T	$\alpha 1 \rightarrow$ peptide	peptide	[24,25]
UDPGalNAc	N-acetyl-galactosaminyl-transferase	GalNAc-T	$\beta \rightarrow 4$	β GlcNAc	[26]
GDPMan	Mannosyl-transferase	Man-T	$\alpha 1 \rightarrow 2$	α/β -mannose	[27]
GDPMan	Mannosyl-transferase	Man-T	$\beta 1 \rightarrow 2$	β GlcNAc	[28]
GDPXyl	Xylosyl-transferase	Xyl-T			

¹ type 1 linkage: Gal $\beta 1 \Rightarrow$ 3GlcNAc; type 2 linkage : Gal $\beta 1 \Rightarrow$ 4GlcNAc

² X refers to a linkage type (3 or 6 or 8)

³ acc, acceptor

⁴ n refers to a number defined by a gene product

The main enzymological properties shared by all mammalian glycosyltransferases are listed in Table 2. While a few glycosyltransferases isolated from natural sources (e.g., ST6Gal, Gal-T1) have been commercially available for many years, other glycosyltransferases have only recently been put on the market as recombinant enzymes. The possibility to express these enzymes in some special cases in prokaryotes and in yeasts may facilitate their future use, since fermentation technology costs are low. This review focuses on the published aspects of expression and production in yeasts and shows that this is a feasible approach to render these enzymes generally available; in some cases however unexpected difficulties may arise.

Basic principles of yeast fermentation

Due to the long and extensive use of yeasts in human nutrition, yeast cultivation has become a field of extensive

Table 2. General properties of mammalian glycosyltransferases

Monomeric, in vivo dimeric?
Type II transmembrane protein ¹
Domain structure consists of cytoplasmic, transmembrane, stem and catalytic domains (see Fig. 1)
Usually N- and/or O-glycosylated, predominantly in the stem region
Ordered bisubstrate mechanism: binds first donor, then acceptor substrate ²
Mn ²⁺ usually cofactor
K _m donor substrate: micromolar range
K _m acceptor substrate: millimolar range
Catalytic activity: usually in the range of 5–20 U/mg

¹ Exceptions have been described

² Rigorous kinetic mechanisms have been determined in few cases only

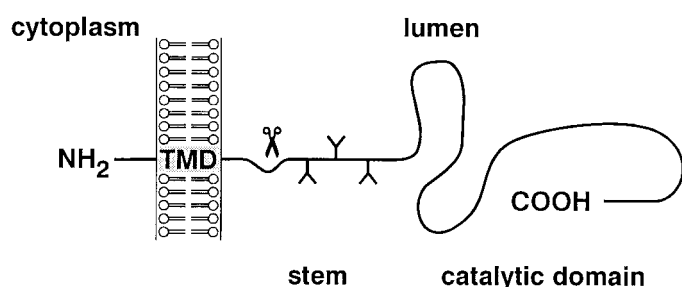


Figure 1. This scheme depicts the domain structure of glycosyltransferases as first described by Paulson and Colley [134]. The scissors mark the approximate position where membrane-bound glycosyltransferase can be converted to secretory forms [135]. TMD designates the transmembrane domain.

study and experience reflected in a vast number of publications. Most importantly, in the context of biosafety those yeasts which are used for the production of human nutrients are considered safe.

The following chapter briefly addresses the main problems. For more detailed information the reader is referred to standard textbooks of biotechnology [31].

Cost-effective production of a heterologous protein in yeasts depends on several factors. First one needs a stable, especially in terms of genetic stability, and reproducible process, fulfilling the legal safety restrictions for genetically modified organisms (GMO). Second, to facilitate downstream processing, a high final concentration of the protein of interest is generally desirable, as well as a high productivity relative to cell number and volume. Because secreted product concentration normally increases with the cell density, high density cultivation systems (80–200 g/l dry mass for *Saccharomyces cerevisiae* on semi-synthetic media) are of general interest.

Following a feasibility study of production in shaker flasks, biotechnologists generally address the question of bioreactor operation mode. Three possibilities, batch, continuous cultivation and fed-batch are available. In a batch cultivation all substrates are given at the start of the fermentation. Therefore, high initial sugar concentrations are often applied. But these are limited by the osmotic sensitivity of the yeasts and by catabolite repression. The possibilities to regulate batch processes are very limited. Once started they can only be influenced by changing one of the continuous parameters like oxygen supply, pH or temperature.

A continuous cultivation allows, in theory, the control of the whole process by manipulation of every component over an unlimited period of time. In practice, the process is often limited by decreasing productivity caused by the selection for non-expressing clones in the system, due to genetic instability. In addition, the product concentration in the harvest is dependent on residence time and normally lower when compared to batch processes. Further difficul-

ties can be caused by attempts to maintain a high oxygen transfer rate (OTR) and sufficient cooling at high dilution rates, which influences the attainable cell density and thus the product concentration.

Fed-batch cultivation allows control of the substrate concentration over a limited period of time by adding nutrients, trace salts and vitamins until the end of the process or upon reaching the maximum volume of the vessel. This high product harvest concentration and the flexibility of the system make it still the method of choice for the expression of heterologous proteins although continuous cultivation approaches are becoming increasingly interesting.

The next step is the choice of an appropriate reactor. Due to the lower doubling time of yeasts (1.5 h) versus approximately 24 h for animal cell lines, monoseptic cultures are generally easier to maintain than for animal cells. The methylotrophic yeasts such as *Pichia pastoris* are additionally protected by the inducer methanol which is toxic or at least growth-inhibiting for a number of other organisms. Therefore, no special sterility technique is necessary. Nevertheless, a number of other requirements have to be fulfilled. During aerobic growth, heat production and oxygen consumption are closely correlated. For each mole of oxygen consumed 455 kJ heat is generated when grown on glucose [32]. The specific oxygen demand of *S. cerevisiae* is 8 mM per h and g of cells, thus somewhat lower than that of *E. coli* (which consumes 10.8). Special care has to be taken for the DO (dissolved oxygen) due to the fact that yeasts have the ability to switch their metabolism to ethanol production at low DO or high glucose concentration [33], thereby changing byproduct formation and growth rate. The maximum attainable cell density can thus be limited by the OTR or the transferable heat, respectively. Therefore, these parameters are of critical importance in the process design, and it seems feasible to keep one of them constant during scale-up. In order to reach high OTR, the presence of a cell wall protecting the plasma membrane in yeasts is of general advantage. It makes the organism comparatively insensitive to shear stress and bubbled aeration, whereas the formation of foam can limit the achievable OTR. If, however, oxygen supply becomes limiting in the process or during scale-up, aeration can be switched to pure oxygen (maximum soluble O_2 increases from 9 to 43 mg/l). In the past years an impressive number of specialized equipment for fermentation has been put on the market. However, the classical stirred vessel most probably due to its flexibility, its availability from lab size to over 100 m³ scale and its comparatively well known scale-up properties

The culture conditions as well as the process design itself is influenced by several parameters including the localization of the protein of interest (intra- vs. extracellularly), the regulation of the promoter (constitutive vs. induced) as well as product stability.

For an intracellular product which is expressed under the control of a constitutive promoter, the process design will

be simply focused on the production of biomass. By contrast, the increasing use of inducible promoters requires at least a two step cultivation: While in the first phase biomass production with high growth rates is generally desirable, the second phase will be adapted for the production of the protein of interest at low growth rates. Expression of secreted products is much preferred due to three main advantages: First, no disruption of the cells is needed; second, the toxicity of the product to the host is of lesser importance; third, down stream processing can be facilitated due to the reduced amount of contaminant proteins. The choice of the carbon source and of the whole media depends on the product market price. For the production of heterologous proteins glucose is commonly used as a carbon source. The temperature is generally set to 28–30 °C and tightly regulated.

Once the process has been successfully established in a small scale bioreactor (typically 1–10 liter), and if higher amounts of product are needed, scale-up will be started due to the fact that production costs are generally inversely related to the scale of fermentation. Depending on the limitations of the process different scale-up criteria are kept constant (e.g., OTR, $k_L a$). In a first step, a one order of magnitude bigger bioreactor is usually investigated.

Since expression of glycosyltransferases in *Pichia pastoris* seems to be of advantage, the following section deals with some special features in using this organism.

The methylotrophic yeast *P. pastoris* can be grown to high density (>100 g/l on glycerol), and has the ability to use methanol as the sole carbon source. It is sold as a kit allowing the gene of interest to be expressed under the control of the strong inducible alcohol oxidase (AOX) promoter. The first step of the biosynthesis is the oxidation of methanol to formaldehyde and hydrogen peroxide, which is subsequently detoxified by catalase while none of the energy of this highly exothermic reaction is harnessed as NAD(P)H or ATP [34]. Additionally three moles of oxygen are required to completely oxidize one mole of methanol, generating 4–6 moles of ATP. Therefore heat production and sufficient oxygen supply are, especially for Mut⁺ strains (which can grow fast on methanol), parameters to be watched during process design and scale-up. Compared to shaker flasks cultivation in a bioreactor is expected to give a 5 to 10 times higher yield, especially for Mut⁺ strains, due to the better oxygen supply [35]. The cultivation process is divided into two phases: In the first phase, biomass is produced using glycerol as a carbon source. Precultures are set up in baffled flasks (5–10% initial reactor volume) and grown to an OD₆₀₀ of 2–6. Subsequently the culture is used to inoculate the reactor containing 4% glycerol, some basic salts and trace elements. DO is set to >20%. If the measured value reaches this critical mark and cannot be maintained by increasing the pressure or by switching to pure oxygen the glycerol feed has to be reduced. The pH is normally set at 5–6 while regulation by the addition of

ammonium hydroxide is recommended. If proteolysis turns out to be critical, the pH can be lowered to 3 during induction. It has also been reported that a higher pH as well as a lower temperature [36] may have the same effect on proteolysis. Another possibility to reduce proteolysis is to include amino acids in the media. Additionally, two strains, SMD 1168 and 1163, deficient in Proteinase A or B have been used to express protease sensitive peptides [37,38]. The addition of antifoam should be evaluated carefully. For Fuc-T III an inhibiting effect of antifoam reducing the activity by 70% has been reported [39]. Interestingly, temperatures above 32°C have been reported to reduce protein secretion. After 18 to 24 h, glycerol will be completely consumed and a density of 90–150 g/l wet cells is expected. In order to reach a higher cell density which is closely linked to a higher product concentration, a glycerol fed-batch step may be applied. When adding glycerol or the inducer methanol, it is recommended to include some trace salt. Siegel and Brierley showed that for an up-scaled process an adapted formulation of trace salts may be favorable [40]. After a density of 50–300 g/l of wet cells is reached methanol is added. The two different methanol utilizing phenotypes, Mut^s and Mut⁺ require different handling: Mut^s strains grow only slowly on methanol thereby consuming lower amounts of methanol and require a reduced methanol feed (1 instead 3.6 ml for Mut⁺ per h and l of initial cultivation volume). Excess methanol (>1–2%) will be toxic for the cells. Accordingly, Mut^s cells have a lower oxygen consumption rate. When the culture is adapted to methanol (normally after 2–4 h) the feed can be doubled. After an additional 2 h the feed can again be increased to 11 for Mut⁺ per h and l and 3 ml/h per l of initial reactor volume for Mut^s strains.

A continuous cultivation process for *Pichia* Mut⁺ has been described: Due to the lower cell density of 100 instead of 120 g dry cells/liter and the lower product concentration of 350 instead of 550 mg/l of product, this approach appears not very promising [41]. Even though a report on tetanus toxin fragment c exists showing an expression level in the 12 g/l scale [42], Fuc-T III, the first documented glycosyltransferase expressed in *P. pastoris*, reached only a concentration of 30 mg/l [39].

Vectors for expression in yeasts

The catalog of plasmid DNA vectors with different markers capable of transforming auxotrophic yeast strains for a variety of cloning purposes has expanded greatly. With few exceptions these plasmids also function as shuttle vectors meaning that they possess bacterial sequences that can be selected for and propagated in *E. coli*.

Two types of vectors are generally used to construct plasmids for yeast transformation: YE_p (yeast episomal plasmid) and YC_p (yeast centromeric plasmid) vectors. Both classes carry a sequence that promotes autonomous

replication in yeast (ARS elements). The ARS elements are yeast origins of DNA replication and their presence promotes high-frequency of transformation. YCp plasmids carry a chromosomal centromere whose sequences are important for plasmid association with the cell's mitotic spindle apparatus. As a consequence the plasmid copy number is limited to 1–2 copies per cell.

The YEp vectors utilize the partitioning system of the endogenous yeast plasmid or episome known as the 2 μ circle to achieve stable high copy propagation [43,44]. The relevant 2 μ sequences function to facilitate equal distribution of plasmid molecules between mother and daughter cells at mitosis. Plasmids harboring 2 μ sequence are maintained at 10–40 copies per cell.

The integrating yeast (YIp) vectors are another type of vector used for expression in yeast. Integration of DNA sequences into the yeast chromosome has the advantage that the yeast can grow in rich rather than selective media to much higher culture density without risking the loss of the desired gene. Therefore, high expression is dependent on the choice of promoters rather than the plasmid copy number. Two different methods have been developed to integrate foreign DNA into the yeast chromosomes: YIp vectors lack an origin for autonomous replication but carry sequences which allow for high-frequency chromosomal integration. These plasmids are linearized by a single restriction cut within the complementary yeast gene on the vector for integrative gene conversion. The insertional integration event results in two copies of the “chromosomal” gene flanking the newly inserted foreign gene. A number of integrating vectors have been used successfully [45,46].

The second approach for gene integration, gene replacement by homologous recombination (integrative gene disruption), is a relatively efficient method for directing DNA to a particular *locus* in the chromosome. The DNA to be integrated is linearized with two restriction cuts and retains flanking sequences matching the desired chromosomal *locus*. The desired gene is engineered into the coding region of the complementary gene for chromosomal insertion, thus disrupting its gene expression thereby substituting expression of the desired gene. This approach has also been used for the elimination of yeast genes that could interfere with efficient expression of the foreign protein [47].

To achieve efficient expression, vectors must contain yeast *promoter* and *terminator* sequences for efficient transcription of the foreign gene as well as *selection markers*. The vector may also contain a *signal sequence* to direct the expression product into the secretory pathway.

Yeast auxotrophic selection markers

The most commonly-used markers for the selection of yeast transformants are *LEU2*, *TRP1*, *URA3*, *HIS3* and *HIS4*; they are used in corresponding mutant strains which are auxotrophic for leucine, tryptophane, uracil and histid-

ine, respectively. Continuous selection requires the use of minimal growth medium lacking the relevant nutrient.

Yeast promoters

For an efficient transcription of a foreign gene the use of yeast promoters was found to be essential [48].

The first promoters used were from genes encoding abundant glycolytic enzymes, eg. alcohol dehydrogenase I (*ADHI*) [48], phosphoglycerate kinase (*PGK*) [49], triose phosphate isomerase (*TPI*) [50,51], or glyceraldehyde-3-phosphate dehydrogenase (*GAP* or *GAPDH*) [52] (Table 3). Glycolytic promoters are the most powerful of *S. cerevisiae*, but they are poorly regulated; this makes them unsuitable for expressing toxic proteins and inappropriate for use in large scale culture where there is a higher risk for the selection of non-expressing cells.

It is preferable to use a tightly-regulated promoter to allow separation of the growth stage from the expression stage. The most powerful tightly-regulated promoters of *S. cerevisiae* are those of the galactose-regulated genes, *GAL1*, *GAL7* and *GAL10* involved in metabolizing galactose. *GAL1*, *GAL7* and *GAL10* mRNAs are rapidly induced >1000-fold to approx 1% total mRNA on addition of galactose [59] (Table 3).

The promoter of the acid phosphatase gene *PHO5* which is regulated by inorganic phosphate concentration has been extensively used for foreign gene expression [54]. The structural features for regulation of the *PHO5* promoter have been studied in details (for review, see [60]).

Another type of promoter belongs to the glucose-repressible promoters. Glucose-repression is a global system regulating the expression of a number of genes including sugar fermentation genes by the availability of glucose. Genes involved in sucrose or galactose metabolism are transcriptionally repressed by glucose. Typical examples of promoters regulated primarily by glucose-repression are those encoding alcohol dehydrogenase II (*ADH2*) [53] or iso-1-cytochrome c (*CYC1*) [55]. The *ADH2* promoter is both powerful and tightly regulated. It is repressed over 100-fold by glucose, thus it can be used for efficient expression of toxic proteins eg. insulin-like growth factor I (IGF-I) [61]. Glucose-repressible systems have a potential serious disadvantage in industrial fermentations: It is difficult to maintain tight glucose-repression under conditions of glucose-limitation which is required to achieve high cell density.

In order to reduce the drawbacks due to certain promoters several hybrid promoters have been designed. The hybrid promoters such as *ADH2/GAPDH* [62] and *GAL10/CYC1* [63], resp., have also been successfully used to express a variety of heterologous gene products. In addition to the previously described promoters other regulated promoter systems have been described, among them the *CUP1* promoter [56]. It is tightly-regulated and inde-

Table 3. Promoters for heterologous gene expression

Promoter	Culture conditions	Expression	References
Alcohol dehydrogenase I (ADH I)	High (2–5%)glucose	constitutive	[48]
Alcohol dehydrogenase II (ADH II)	Low (0.1–0.2%)glucose	inducible	[53]
Phosphoglycerate kinase (PGK)	High (2–5%)glucose	constitutive	[49]
Triose phosphate isomerase (TPI)	High (2–5%)glucose	constitutive	[50,51]
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	High (2–5%)glucose	constitutive	[52]
Acid phosphatase (PHO5)	Phosphate-deficient medium	inducible	[54]
Cytochrome c1 (CYC1)	Glucose as carbon source	repressible	[55]
Metallothionein (CUP1)	Copper (0.03–0.1 mM)	inducible	[56]
Alcohol oxidase (AOX)	Methanol	inducible	[57]
Methanol oxidase (MOX)	Methanol	inducible	[58]
Galactose-regulated genes (GAL1, GAL7, GAL10)	Galactose	inducible	[59]

pendent of culture parameters. The concentration of Cu^{2+} ions for induction depends on the copper-resistance of the host strain from 0.01 mM (no CUP1 gene) to 0.5 mM (>6 gene copies).

The expression of heterologous proteins in yeasts other than *S. cerevisiae* has become more popular over the past few years particularly in the methylotrophic yeast *Pichia pastoris* and *Hansenula polymorpha*. Alcohol oxidase (AOX), the first enzyme in the methanol utilization pathway in *P. pastoris*, is dramatically induced in cells grown on methanol as the sole carbon source [57]. A similar physiologic response has been reported for the *H. polymorpha* methanol oxidase (MOX) [58]. By using recombinant plasmids containing the AOX or MOX gene promoters or the promoter for the formate dehydrogenase (*FMDH*) gene [64], it has been possible to produce a number of heterologous gene products to high levels in the methylotrophic yeasts [for a general review see 35].

Signal sequences

Heterologous protein may be secreted from yeasts using either a foreign signal (often derived from the protein being secreted) or a yeast signal. Since signal sequences are recognized with low specificities in yeast [65], it could be assumed that foreign signals would work as efficiently as those from yeast but this is often not the case. Early attempts to secrete foreign proteins from *S. cerevisiae* utilized the protein's own signal [66–68]: the expression levels were very often very low with only a proportion of the protein being secreted. Thus the use of foreign leaders often results in intracellular accumulation. Therefore, for most cases of heterologous protein secretion from yeast it is preferable to use a yeast signal sequence. Much work has been carried out using homologous *S. cerevisiae* signal sequences. The

most widely used are the signal peptides from acid phosphatase [69], invertase [51,70] and α factor [71–73]. Many foreign proteins have now been secreted from yeast using the α -factor leader and this system has been demonstrated to be generally applicable.

Yeast terminators

Yeast transcriptional terminators are usually present in expression vectors for efficient mRNA 3' end formation. Terminators of prokaryotic or higher eukaryotic genes are normally not active in yeasts though there are some exceptions such as the *Drosophila* ADE8 gene [74]. Efficient termination is required for maximal expression; indeed Zaret and Sherman [75] demonstrated that deletion of 'termination' sequences 3' of the CYC1 gene resulted in longer mRNA and a dramatic reduction in mRNA level.

In order to simplify the vector's construction, the yeast terminator corresponding to the yeast promoter is usually used (eg. *TRP1*, *ADHI*, *PHO5* or *GAP*); another alternative is the use of a terminator from 2 μ circle, eg. FLP [67] or D gene terminator [76].

Comparison of prokaryotic and eukaryotic cell expression

The choice of an expression system for the high-level production of recombinant proteins depends on many factors. These include scientific biological criteria such as cell growth characteristics, expression levels, intracellular and extracellular expression, posttranslational modifications and biological activity of the protein of interest as well as its intended use. In addition, commercial patent-related criteria also play a role in choosing the appropriate expression system. The relative advantages of the yeast expres-

sion system as compared to bacterial or mammalian cell expression systems will be briefly described in this paragraph; for a more exhaustive treatment see the review of Marino [77].

In comparison with yeasts or mammalian cells, the major drawbacks of expression in *E. coli* are the inability to perform many of the posttranslational modifications found in eukaryotic proteins, the lack of a secretion mechanism for the efficient release of protein into the culture medium, the limited ability to facilitate extensive disulfide bond formation and the production of inclusion bodies (for a general review, see [78]). Many eukaryotic proteins retain their full biological activity in a nonglycosylated form; therefore, they can be produced in *E. coli*; for those which require glycosylation or for those whose complex tertiary structure depends in part on disulfide bond formation, yeast and mammalian cell expression systems are necessary. Indeed, a particular advantage of these two systems is that the foreign protein may be directed into the secretory pathway usually by fusion of the mature form of the recombinant protein to a given peptide signal. Along the secretory pathway protein folding, disulfide bond formation and glycosylation take place. Secretion of properly folded proteins which is crucial for full biological activity, is one of the major factors determining the choice of yeasts or mammalian cells as hosts for heterologous protein expression. This is due to the fact that the direct capture of active product from conditioned medium eliminates the need for costly low-yielding cell-disruption or refolding steps [79].

Glycosylation is both the most common and the most complex form of posttranslational modification [80]. The majority of therapeutically relevant proteins are glycosylated in their natural forms and should also be glycosylated as recombinant proteins in order to get the correct biological activity. Thus, monitoring of glycosylation pattern in quality control of recombinant proteins to assure product safety, efficiency and consistency has become increasingly important. Whereas *N*-glycosylation pathways associated with the endoplasmic reticulum are highly conserved between yeast and mammalian cells, chain elongation and termination occurring in the Golgi apparatus are different as outlined below. Accordingly, yeast cells recognize the same *N*-glycosylation amino acid sequence as higher eukaryotic cells. The glycosyl groups on yeast glycoproteins consist primarily of mannose residues appended in different linkages to the core glycosyl units (for review, see [27]). Since the recombinant glycoproteins generated in *S. cerevisiae* are of the high-mannose type, they will be recognized by mannose receptors on various cells and removed when injected into the circulation of mammalian species. In addition, non-human glycosylation patterns are potentially immunoreactive. In fact, *S. cerevisiae* is known to synthesize large polymannans consisting of 50 to 100 mannose residues, a phenomenon also referred to as hyperglycosylation. The bulkiness of these glycans may considerably impair

biological activity of the recombinant protein and, therefore, negate any advantage of the microbial eukaryote *S. cerevisiae* expression system over *E. coli* or mammalian cells. Mannan mutants (*mnn*) have been isolated, which synthesize shorter mannan chains but do not grow as well as other yeast strains. Other yeast species such as *P. pastoris* and *H. polymorpha* seem less prone to hyperglycosylating heterologous proteins [40]. The average mannose chain length produced in the latter two yeasts is only 8–14 monomers which is comparable with complex type glycans. In fact, in all cases of recombinant proteins with intrinsic commercial value for use as catalysts or as model compounds to rationally design biomimetics, methylotrophic yeasts are now the preferred option as expression systems [57].

Because of the hyperglycosylation occurring in some yeast expression systems, in many cases mammalian cells are the preferred host cells for the generation of recombinant glycoprotein therapeutics. For these aspects the reader is referred to [81]. In the special case of recombinant glycosyltransferases an additional drawback of higher eukaryotic cell hosts is the observation that these cells may express silenced glycosyltransferases genes upon transfection [82,83]; this may lead to confusion in interpreting the specificities of recombinant glycosyltransferases. Moreover, animal cells are known to release a number of endogenous Golgi-associated glycosyltransferases during growth [84,85] which may obscure the specificity of the recombinant enzyme to some extent. Such a problem does not exist when using *E. coli* or yeasts as expression systems because of their limited repertoire of glycosyltransferases which is restricted to mannosyl-, glucosyl- and core GlcNAc-transferases [86].

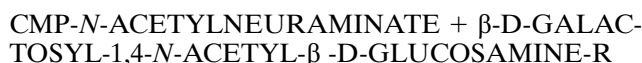
In summary, the selection of a particular expression system from *E. coli*, yeasts or mammalian cells depends on the nature and use of the recombinant protein and the related production costs. The yeast expression system combines the ease, simplicity and low costs of bacterial expression systems: Like bacteria, yeasts are simple to cultivate on inexpensive growth media. The reader is referred to Datar *et al.* [87] who have analyzed the economic issues associated with protein production in bacterial and mammalian cells: They concluded that for each recombinant protein, it is necessary to evaluate the production process as a whole. Since each protein has its own requirements in terms of folding, glycosylation and maturation-associated cleavage or other posttranslational modifications, extensive pilot-scale studies are essential for rigorous comparative evaluations.

Specific section

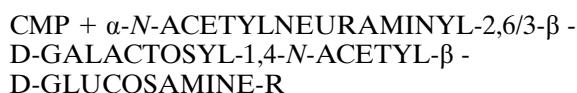
Sialyltransferases

Sialyltransferases are glycosyltransferases that transfer sialic acid from the donor substrate CMP-sialic acid to

different types of oligosaccharides according to the following general reaction:



→



As shown in Table 1 the sialyltransferases are grouped according to the linkage type they catalyze. Within the family of sialyltransferases and in contrast to other families of glycosyltransferases, practically no sequence homology was observed except for two conserved motifs, the “L-sialylmotif” and the “S-sialylmotif” located in the catalytic domain. Both of them have been shown by Datta and Paulson [88,89] to be involved in the binding of the donor substrate.

The yeast expression data are compiled on Table 4.

α -2,3-Sialyltransferases

To date four different α -2,3-sialyltransferases have been cloned. Their *in vivo* substrate specificities have been extensively studied by Tsuji et al. [16]. Only one out of these four has yet been expressed in yeasts, the ST3GalIII (GenBank Accession # m97754). It has the following specificities for the acceptor substrates, in the order of decreasing rate of the reaction: Gal β 1,3GlcNAc > Gal β 1,4GlcNAc > Gal β 1,3GalNAc.

A truncated form (aa 29-374) of the rat ST3GalIII [94] has been expressed in *S. cerevisiae* H23 and H626 [90]. The enzyme was *N*-terminally fused to the hsp150 Δ polypeptide, a carrier which has previously been used for heterologous proteins to confer proper folding and secretion into the growth media [95]. The enzyme was shown to be active but remained intracellularly. Mattila et al. [90] showed that the protein was poly-mannosylated on one or both of its

potential *N*-glycosylation sites indicating that the enzyme was transported at least to the Golgi apparatus. Interestingly, tunicamycin treatment abolished activity completely suggesting that glycosylation of ST3GalIII is required for proper folding and activity. Since the enzyme remained intercalated in the cell wall whole live yeast cells were used as enzymatic catalyst for the synthesis of sialyl α 2,3-*N*-acetylactosamine a prerequisite for the sialyl Lewis^x epitope synthesis. Thus 110 mU/l were obtained while the same enzyme expressed in insect cells has been shown to yield 25 U/l after 72 h of incubation [96].

α -2,6-Sialyltransferases

The family of α -2,6-sialyltransferases can further be subdivided by the acceptor used. One subgroup acts on galactose while the other incorporates sialic acid on *N*-acetylglactosamine. To date, only one has been expressed in yeast. Krezdorn et al. [91] expressed a full-length form of human α -2,6-ST6Gal (GenBank Accession # X17247 EC 24991) which incorporates sialic acid on Gal β 1,4GlcNAc. The first host to be used was a protease-deficient strain of *S. cerevisiae* BT₁₅₀ which later was cultivated and up-scaled by Borsig et al. [92] to the 150 l scale. It yielded an activity of 0.3 U/l in the yeast lyophilisate. The ST6Gal was produced in a glycosylated form containing two oligomannose units as shown by Endoglycosidase-H treatment. As concluded by pulse chase analysis, it was located in the endoplasmic reticulum (ER) or an early Golgi compartment. It appeared to be translocated into the endoplasmic reticulum by a posttranslational mechanism as not uncommonly found in this host (for a review, see [97]). The recombinant ST6Gal had similar Michaelis constants (101 μ M) as the native rat enzyme (158 μ M) for the donor substrate. The constants for the acceptor substrates however were markedly different. For the recombinant form a K_m of 4.65 mM was determined while the native rat enzyme revealed a K_m of 1.67 mM. Subsequently, this recombinant ST has been used by van Dorst [98] to explore the substrate specificities of ST6Gal.

Table 4. Expression of sialyltransferases in yeast

Enzyme	Form	Vector	Host cell (mU/mg Protein)	Specific activity	Volume specific activity (mU/l)	localization	reference
α 2,3 ST	soluble hsp 150 Δ fusion-protein	pKTH4636	<i>S. cerevisiae</i> H23, <i>S. cerevisiae</i> H626	not reported	110	cell wall	[90]
α 2,6 ST	full-length soluble	pDP34 pDP34	<i>S. cerevisiae</i> BT ₁₅₀ <i>S. cerevisiae</i> BT ₁₅₀	0.3	300 1	ER, early Golgi cell	[91,92] Borsig et al., unpublished
α 2,6 ST	soluble	pPIC9	<i>P. pastoris</i> KM71	0.086	300	media	[93]

To facilitate downstream processing, attempts were made at expressing ST6Gal as a soluble form to release it into the supernatant. A truncated form lacking the transmembrane domain has thus been expressed in *S. cerevisiae* BT₁₅₀: The enzyme was weakly active (1 mU/l). The reason for this low activity was not further explored but might have been due to hyperglycosylation of the *N*-glycosylation site located close to the catalytic portion [92]. The same enzyme was therefore expressed in *P. pastoris* yielding a volumetric activity 0.3 U/l (manuscript in preparation).

Galactosyltransferases

β -1,4-Galactosyltransferase

β -1,4-galactosyltransferase (gal-T1) (EC 2.4.1.38 GenBank Accession M22921) is one of the best known glycosyltransferases. Its general reaction is

UDP-GALACTOSE + *N*-ACETYL- β -D-GLUCOSAMINYLGlycopeptide

→

UDP + β -D-GALACTOSYL-1,4-*N*-ACETYL- β -D-GLUCOSAMINYLGlycopeptide

The enzyme also catalyzes the transfer of galactose to glucose in the presence of α -lactalbumin (E.C. 2.4.1.22) and many other glycoconjugates exhibiting terminal β -GlcNAc-residues. Gal-T1 (see Table 1) is a Golgi-associated glycosyltransferase located in *trans* Golgi cisternae and possibly on cell surfaces. The enzyme is a type II membrane protein consisting of a single polypeptide. Gal-T1 is solubilized *in vivo* and released from the cells in an enzymatically truncated active form [85]. The peptide contains a single *N*-glycosylation site and various *O*-glycosylation sites [99].

Kreuzdorn et al. [100] were the first to report heterologous expression of full length gal-T1 using a yeast expression system (Table 5). The vector used contained the *PHO5* promoter, the cDNA encoding full-length gal-T1 and the *PHO5* terminator. The *PHO5* promoter fragment lacks the upstream regulatory fragments and therefore acts like a constitutive promoter. Recombinant gal-T1 produced in the *S. cerevisiae* BT₁₅₀ strain was investigated by metabolic labeling followed by immunoprecipitation and shown to be retained in the endoplasmic reticulum [91]. Kreuzdorn et al. [100] assumed that the signal for Golgi targeting retention was not recognized in yeasts and this was further supported by the results obtained with an other glycosyltransferase, the ST6Gal [91]. The transmembrane anchor which acts as the secretion leader for these type II membrane proteins, appears to determine the expression level of gal-T1 which was found to be much lower than the expression level of gal-T1 fused to the membrane anchor of the α 1,2mannosyl-

transferase (Mnt1p) despite the fact that all the transformants produce equal mRNA level; this may be due to the fact that the transmembrane domain of gal-T1 contains amino acids which are rarely used in yeast genes [101]. In addition, the membrane anchor of the Mnt1p was able to target gal-T1 to the yeast Golgi complex [102]. Altogether, this could also explain the high difference observed between the expression level obtained by Kreuzdorn et al. [100] (gal-T1 represents 0.01% of crude extract protein) and Schwientek and Ernst [101] (gal-T1 represents 0.15% of crude extract protein).

Since the handling of a membrane protein for purification and use *in vitro* for organic synthesis is difficult, attempts were made to express soluble forms of gal-T1 [101,104]. In both cases, the transmembrane domain was replaced by a yeast signal sequence which allows translocation to the secretory pathway. When the invertase signal sequence was used, gal-T1 was produced as a soluble enzyme retained in the yeast cell [104]; when the α -factor signal sequence was used, 62% of the produced soluble gal-T1 was secreted into the culture medium [101]. This major difference in the protein localization was probably due to the signal peptide.

On the laboratory scale Kleene et al. [104] obtained an expression level of 200 mU/l culture. By using fermentation in fed batch technique it was possible to increase the expression level of recombinant soluble gal-T1 from 200 mU/l to 705 mU/l [105]; this study also demonstrated for the first time that heterologous expression of a glycosyltransferase is possible on a large scale (use 150 l fermentor). This offers a good alternative to natural sources of gal-T1 like human milk or bovine colostrum.

The recombinant soluble gal-T1 expressed in *S. cerevisiae* BT₁₅₀ was purified to homogeneity through three successive affinity chromatographies and was shown to be *N*-hyperglycosylated [103]. The *N*-hyperglycosylation had no impact on gal-T1 activity since it was observed that the specific activity of the purified recombinant gal-T1 was on the same order as the catalytic activity of the human gal-T1. The *N*-hyperglycosylated biochemically pure recombinant gal-T1 was shown to be very heterogeneous when analyzed by immunoblotting. In order to obtain a more homogeneous protein, the unique *N*-glycosylation site was removed by using site directed-mutagenesis; the soluble mutagenized recombinant gal-T1 expressed in *S. cerevisiae* BT₁₅₀ was scaled up to 60 U in a 150 l fermentor which represents an expression level of 400 mU/l. After purification the specific activity of the soluble mutated recombinant gal-T1 was 6.9 U/mg and the *K_m* values were comparable to those reported for the human gal-T1.

Aiming at simplifying purification of recombinant gal-T1 expressed in *S. cerevisiae*, Borsig et al. [106] introduced an His₆-tag to the *N*-terminus of gal-T1 (his-gal-T1). Binding efficiency of his-gal-T1 was found to be impaired by the bulky *N*-glycan previously described by Malissard et al.

Table 5. Expression of gal-T1 in yeasts

Enzyme	Form	Vector	Host cell	Specific activity (mU/mg protein)	Volume specific activity (mU/l)	Localization	Reference
gal-T1	Full length	pDP34	<i>S. cerevisiae</i> BT ₁₅₀	0.42	—	intracellular	[100]
		pGTraI	<i>S. cerevisiae</i> BJ1991		level of detection	intracellular	[101]
	Soluble	pDP34	<i>S. cerevisiae</i> BT ₁₅₀	0.66	705 (fermentation)	intracellular	[105]
		pMFGTI	<i>S. cerevisiae</i> BJ1991	—	—	secreted	[101]
		pPIC9	<i>P. pastoris</i> KM71	15	400 (lab scale)	secreted	Malissard et al., unpublished
	Soluble fused to N-terminus part of Mnt1 soluble <i>N</i> -deglycosylated	phMGTI	<i>S. cerevisiae</i> BJ1991	1.47	—	intracellular	[101]
		pDP34	<i>S. cerevisiae</i> BT ₁₅₀		220 (lab scale)	intracellular	[103]
					425 (fermentation)	intracellular	[103]
					(fermentation)	intracellular	[136]
		pPIC9	<i>P. pastoris</i> KM71	0.17	350 (lab scale)	intracellular	[Malissard et al., unpublished]
	His-tagged soluble	pDP34His	<i>S. cerevisiae</i> BT ₁₅₀	—	25 (lab scale)	intracellular	[106]
	His-tagged soluble <i>N</i> -deglycosylated	pDP34His	<i>S. cerevisiae</i> BT ₁₅₀	—	15.6 (lab scale)	intracellular	[106]

[103] close to the His-tag. Removal of the unique *N*-glycosylation site restored binding of his-gal-T1 to the Ni-NTA resin.

Recently, gal-T1 was also expressed in *P. pastoris*. Briefly, gal-T1 was found to be secreted and active (400 mU/l on the laboratory level) while the non-*N*-glycosylated form was active but not secreted (Malissard *et al.* manuscript in preparation). These data suggest that *N*-glycosylation of gal-T1 is required for efficient secretion by *P. pastoris* although it is not required for enzyme activity. Such results were already reported for proteins expressed in *S. cerevisiae*; indeed the secretion of acid phosphatase [107], *Mucor pulsus* rennin [108] and an analog of human tissue plasminogen activator [109] are inhibited when *N*-glycosylation sites are removed. For expression in *P. pastoris* there is no general rule concerning the need of *N*-glycosylation for efficient secretion. Indeed, secretion efficiency of α -*N*-acetylgalactosaminidase is strongly decreased when the *N*-glycosylation is missing [110], but Tsujikawa *et al.* [111] reported secretion of a variant of human single-chain urokinase-type plasminogen activator without an *N*-glycosylation site. The gal-T1 produced by *P. pastoris* is recognized directly by polyclonal antibodies and when analyzed on SDS-PAGE it comigrates with the human gal-T1 indicating that no hyperglycosylation takes place (Malissard *et al.* unpublished results). The secretion of gal-T1 in the culture medium represents a major advance with respect to production of enriched preparations.

In recent years, interest for gal-T1 and for glycosyltransferases in general has been aroused by their potential usefulness as tools for the synthesis of oligosaccharides *in vitro* (for reviews, [112–114]), for the remodeling of glycan chains of natural or recombinant proteins and in case of gal-T1 for the enzymic galactosylation of non-natural glycosides.

Early uses of natural gal-T1 in synthesis have been reported by Barker *et al.* [115] who incorporated Gal into GlcNAc immobilized on Sepharose. More recently, Herrmann *et al.* [116] reported the use of transformed whole yeast cells carrying a plasmid for the heterologous expression of a soluble human gal-T1 for the synthesis of *N*-acetyllactosamine.

Above we gave two examples where gal-T1 was used for the synthesis of disaccharides, but it can also be used in combination with organic synthesis or with other glycosyltransferases for the synthesis of more complex oligosaccharides [117]. In 1997, Unverzagt [118] described an effective use of gal-T1 and ST6Gal to complete a chemical synthesis which was difficult to achieve by chemical means. A heptasaccharide asparagine conjugate was galactosylated and sialylated in the presence of alkaline phosphatase to afford a sialylated undecasaccharide in 86% yield. There are many other examples where gal-T1 in combination with other glycosyltransferases has been used for the synthesis of large glycan structures [119–122].

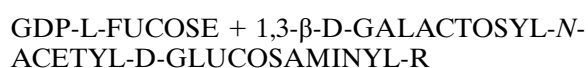
Gal-T1 can also be used for remodeling of glycan chains of natural or recombinant glycoproteins. Witte and coworkers [123] reported the synthesis of an unnatural glycoform of ribonuclease. Ribonuclease B contains a single *N*-linked glycosylation site and exists as a series of high-mannose glycoforms. Treatment with endoglycosidase H gives the ribonuclease derivative, GlcNAc-RNase, with a single GlcNAc attached at this site. Treatment with gal-T1 and a cofactor recycling system gives rise to the disaccharide-linked protein in 76% yield. Further treatment with α -1,3-fucosyltransferase (Fuc-T) or α -2,3-sialyltransferase (ST) or ST followed by Fuc-T gives rise to the predicted glycoprotein products in yields of 72%, 85% and 74% (for the Fuc-T-catalyzed step), respectively. The product of the three-step synthesis is a protein-bound form of sialyl Lewis x. Other examples can be found in Ju and Kean [124] and Schneider *et al.* [125].

Fucosyltransferases

Among the fucosyltransferases indicated in table 1 only Fuc-TIII and Fuc-TV have been described as recombinant enzymes expressed in lower eukaryotes.

Fucosyltransferase III

Fuc-TIII (Galactoside 3(4)-L-fucosyltransferase E.C. 2.4.1.65 (Genbank accession X87810) also known as blood group Lewis alpha-4-fucosyltransferase catalyzes the following reaction:



→



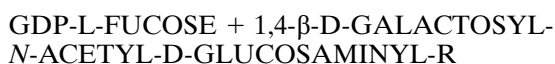
This enzyme has been expressed in *P. pastoris* [34] as a truncated form to facilitate release to the supernatant. In fact, it was observed that the enzyme remained cell-associated for 30 h of continuous fermentation before being released. The enzyme present in the supernatant was detected as a 43 kDa band by SDS-gel electrophoresis and immunoblotting and shown to account for 10% of the proteins in the supernatant. The yield was 113 U per liter. The enzymic properties were characterized using a variety of small molecular weight acceptors. As predicted, the enzyme incorporated fucose into type 1¹ structures and at a much lower rate into type 2 structures. In fact, LacNAc was not an acceptor at all. As already shown by Herrmann *et al.* for

¹the linkage types are defined in Table 1.

β 1,4gal-T1 [116], the cell-associated enzyme could be used as an immobilized (reusable) catalyst to synthesize a selectin ligand [39].

Fucosyltransferase V

Fuc-TV has been expressed in the filamentous fungus *Aspergillus niger* var. *Awamori*. This type of fucosyltransferase belongs to a gene family which has been described in humans [126]. The physiological significance and expression of this enzyme have not been elucidated except for a report describing its expression in pancreatic cancer [127]. Its general reaction is as follows:



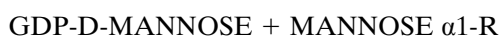
→



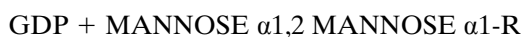
The acceptor may contain a terminal α 2,3 sialic acid; in this case the k_{cat} is 6.5 times lower. The enzyme has been expressed as a fusion protein linked to glucoamylase by a kex 2 proteolytic site. The enzyme was then truncated to a soluble catalytically active form whose cumulative activity reached 300 U/l in the supernatant. Downstream processing involved a 20–60% ammonium sulfate cut and chromatography on phenyl-Sepharose [128]. Besides extensive kinetic characterization an inhibitor was developed [129] showing the usefulness of recombinant enzymes in almost unlimited supply.

α -1,2-Mannosyltransferase

The best characterized mannosyltransferases are those in *S. cerevisiae* which harbour a family of related genes known as KRE2/MNT1. The general reaction is as follows:



→

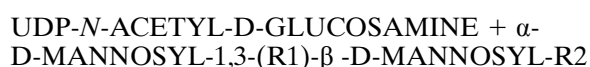


For standard assays α -methyl-mannose has been used. Since in yeasts the α 1,2 linked mannose residues occur in the core *N*-glycan, in the outer chains, as well as in *O*-linked glycans, delineation of the specificity of a gene product with mannosyltransferase activity is necessary to assign the physiological function. For this purpose two mannosyltransferase gene products from *S. cerevisiae*, i.e., Ktr1p, Kre2p/Mnt1p, have been expressed in *P. pastoris* and their enzymic properties compared.

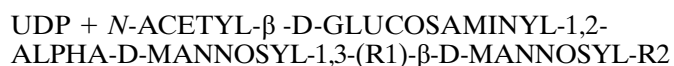
Both enzymes were expressed as truncated enzymes without cytoplasmic and transmembrane domains fused with the cleavable signal sequence of the PHO1 gene product. The released enzymes were identified by SDS-PAGE and immunoblotting and shown to be very pure. The soluble form of recombinant Ktr1p amounted to 400 mg/l while Kre2p was 40 mg/l. Substrate specificity studies then showed that both enzymes can utilize *N*-type glycans [130].

Cell engineering

Heterologous expression of human glycoproteins in yeasts has been on the agenda of several companies since the early eighties. Soon it became obvious that the glycosylation pattern of yeast consisting of the polymannose outer chains differs from the complex *N*-glycans present in human glycoproteins (for review, see [131]). Moreover, the *O*-glycosidic glycan chains are entirely different. Importantly, the mechanism of core *N*-glycosylation in human cells is highly conserved in many respects and indistinguishable from the mechanisms in yeast. The divergence of the *N*-glycosylation pathway between higher and lower eukaryotes occurs in the Golgi apparatus where α mannosidase I trims the oligomannose structure to Man₅GlcNAc₂. *N*-acetylglucosaminyltransferase I (EC 2.4.1.101) thereafter elongates this core oligomannose structure according to the following reaction:



→



This enzyme is absent in yeast as well as enzymes further along the pathway involved in structuring a “human type” biantennary complex *N*-glycan.

An ambitious endeavour which has been initiated in several laboratories aims at introducing the complete metabolic armamentarium into yeast to enable them to structure complex *N*-glycans. The strategy takes advantage of several glycosylation mutants such as Δochl , Δmnnl and Δmnn4 [132] to prevent polymannosylation. However, in order to synthesize the substrate for *N*-acetylglucosaminyltransferase I an α -mannosidase I-like enzyme must cleave the α 1,2 linked mannose residues. This has recently been rendered possible by expressing the α 1,2mannosidase from *Aspergillus saitoi* in yeast cells. Elegantly, to have this enzyme working at the proper site of the secretory pathway, a Golgi to ER retrieval signal (the HDEL sequence) was linked to the C-terminus [132]. The consequence of the introduction of this enzyme to *N*-glycosylation was examined on carboxypeptidase Y and on cell wall mannoprote-

ins. Trimming to the Man₅ species was found to be approx 25% and 10%, respectively. To achieve elongation and termination of complex glycans, additional glycosyltransferases and nucleotide sugar transporters are required. Surprisingly, heterologous expression of human β 1,4galactosyltransferase led to the incorporation of galactose into several yeast proteins [103], suggesting the presence of the metabolic pathways to synthesize UDP-gal and to transport it across the Golgi membranes. This has in fact been confirmed recently by a direct demonstration of a UDP-gal/UMP antiporter in *S. cerevisiae* Golgi membranes [133].

The accomplishment of the goal to develop a *S. cerevisiae* strain that produces "humanized" heterologous glycoproteins appears still remote. However, the endeavour may spin off more interesting and unexpected results.

Acknowledgments

Part of the experimental work from this laboratory cited in this work was supported by SNF-SPP-Biotech grant 5002-46084 to EGB and EU-grant BIO4-CT95-0138.

References

- Varki A (1993) *Glycobiology* **3**: 97–130.
- Yarema KJ, Bertozzi CR (1998) *Current Opinion in Chemical Biology* **2**: 49–61.
- McEver RP (1997) *Glycoconj J* **14**: 585–91.
- Hayden FG, Osterhaus AD, Treanor JJ, Fleming DM, Aoki FY, Nicholson KG, Bohnen AM, Hirst HM, Keene O, Wightman K (1997) *N Engl J Med* **337**: 874–80.
- Stahl PD, Ezekowitz RAB (1998) *Current Opinion in Immunology* **10**: 50–5.
- Bach FH (1998) *Annu Rev Med* **49**: 301–10.
- Bowman KG, Hemmerich S, Bhakta S, Singer MS, Bistrup A, Rosen SD, Bertozzi CR (1998) *Chem & Biol* **5**: 447–60.
- Cooper DK, Koren E, Oriol R (1994) *Immunol Rev* **141**: 31–58.
- Winn R, Vedder N, Ramamoorthy C, Sharar S, Harlan J (1998) *Blood Coagul Fibrinolysis* **9 Suppl 2**: S17–23.
- Briggs JB, Oda Y, Gilbert JH, Schaefer ME, Macher BA (1995) *Glycobiology* **5**: 583–88.
- Garegg PJ (1997) *Adv Carbohydr Chem Biochem* **52**: 179–205.
- Sadler JE, Beyer TA, Oppenheimer CL, Paulson JC, Prieels J-P, Rearick JJ, Hill RL (1982) *Meth Enzymol* **83**: 458–514.
- Kleene R, Berger EG (1993) *Biochim Biophys Acta* **1154**: 283–325.
- Schachter H (1994) In *Molecular Glycobiology* (Fukuda M, Hindsgaul O, eds) pp 88–162. Oxford: IRL Press.
- Tsuji S (1996) *J Biochem* **120**: 1–13.
- Tsuji S, Datta AK, Paulson JC (1996) *Glycobiology* **6**: R 5–7.
- Breton C, Bettler E, Joziassé DH, Geremia RA, Imberty A (1998) *J Biochem* **123**: 1000–9.
- Almeida R, Amado M, David A, Levery SB, Holmes EH, Merckx G, Vankessel AG, Rygaard E, Hassan H, Bennett EP, Clausen H (1997) *J Biol Chem* **272**: 31979–91.
- Hennet T, Dinter A, Kuhnert P, Mattu TS, Rudd PM, Berger EG (1998) *J Biol Chem* **273**: 58–65.
- Amado M, Almeida R, Carneiro F, Levery SB, Holmes EH, Nomoto M, Hollingsworth MA, Hassan H, Schwientek T, Nielsen PA, Bennett EP, Clausen H (1998) *J Biol Chem* **273**: 12770–78.
- Mollicone R, Candelier JJ, Reguigne I, Couillin P, Fletcher A, Oriol R (1994) *Transfus Clin Biol* **1**: 91–7.
- Lowe JB (1997) *Kidney International* **51**: 1418–26.
- Taniguchi N, Ihara Y (1995) *Glycoconjugate J* **12**: 733–38.
- Bennett EP, Weghuis DO, Merckx G, Vankessel AG, Eiberg H, Clausen H (1998) *Glycobiology* **8**: 547–55.
- Clausen H, Bennett EP (1996) *Glycobiology* **6**: 635–46.
- Furukawa K, Soejima H, Niikawa N, Shiku H, Furukawa K (1996) *J Biol Chem* **271**: 20836–844.
- Burda P, Aebi M (1999) *Biochim Biophys Acta* **1426**: 239–257.
- Watt GM, Revers L, Webberley MC, Wilson IB, Flitsch SL (1997) *Carbohydr Res* **305**: 533–41.
- Baisch G, Oehrlein R, Ernst B (1996) *Bioorg Med Chem Lett* **6**: 749–54.
- Baisch G, Oehrlein R, Katopodis A, Ernst B (1996) *Bioorg Med Chem Lett* **6**: 759–62.
- Scragg AH (1991) In *Biochem Biotechnology: Ellis Horwood Series* New York NY: Horwood.
- Bonnet JA, de Kok HE, Roels JA (1980) *Antonie Van Leeuwenhoek* **46**: 565–76.
- Fiechter A, Fuhrmann GF, Käppli O (1981) *Adv Microb Physiol* **22**: 123–83.
- Anthony C (1982) *The Biochemistry of Methylotrophs*. New York: Academic Press, pp 269–95.
- Romanos MA (1995) *Current Opinion Biotechnol* **6**: 527–33.
- Tottrup HV, Carlsen S (1990) *Biotechnol Bioengineering* **35**: 339–48.
- Talmont F, Sidobre S, Demange P, Milon A, Emorine LJ (1996) *FEBS Lett* **394**: 268–72.
- Weiss HM, Haase W, Michel H, Reiländer H (1995) *FEBS Lett* **377**: 451–56.
- Gallet PF, Vaujour H, Petit JM, Maftah A, Oulmouden A, Oriol R, Lenarvor C, Guilloton M, Julien R (1998) *Glycobiology* **8**: 919–25.
- Siegel RS, Brierley RA (1989) *Biotechnol Bioengineering* **34**: 403–4.
- Digan ME, Lair SV, Brierley RA, Siegel RS, Williams ME, Ellis SB, Kellaris PA, Provow SA, Craig WS, Velicelebi G, Arpold MM, Thill GP (1989) *Bio/Technology* **7**: 160–64.
- Clare JJ, Rayment FB, Ballantine SP, Sreekrishna K, Romanos MA (1991) *Bio/Technology* **9**: 455–60.
- Rose AB, Broach JR (1990) *Meth Enzymol* **185**: 234–79.
- Schneider JC, Guarente L (1991) *Meth Enzymol* **194**: 373–88.
- Moracci M, La Volpe A, Pulitzer JF, Rossi M, Ciaramella M (1992) *J Bacteriol* **174**: 873–82.
- Smith RA, Duncan MJ, Moir DT (1985) *Science* **229**: 1219–24.
- Rothstein R (1991) *Meth Enzymol* **194**: 281–301.
- Hitzeman RA, Hagie FE, Levine HL, Goeddel DV, Ammerer G, Hall BD (1981) *Nature* **293**: 717–22.
- Tuite MF, Dobson MJ, Roberts NA, King RM, Burke DC, Kingsman SM, Kingsman AJ (1982) *EMBO J* **1**: 603–8.
- Moir DT, Davidow LS (1991) *Meth Enzymol* **194**: 491–507.
- Melnick LM, Turner BG, Puma P, Price-Tillotson B, Salvato KA, Dumais DR, Moir DT, Broeze RJ, Avgerinos GC (1990) *J Biol Chem* **265**: 801–7.

- 52 Holland JP, Holland MJ (1980) *J Biol Chem* **255**: 2596–605.
- 53 Price VL, Taylor WE, Clevenger W, Worthington M, Young ET (1990) *Meth Enzymol* **185**: 308–18.
- 54 Hinnen A, Meyhack B, Heim J (1989) In *Yeast Genetic Engineering* (Brake PJ, and Valenzuela P, eds) Butterworths, pp 193–213.
- 55 Guarente L, Ptashne M (1981) *Proc Natl Acad Sci USA* **78**: 2199–203.
- 56 Etcheverry T (1990) *Meth Enzymol* **185**: 319–29.
- 57 Sudbery PE, Gleeson MA, Veale RA, Ledeboer AM, Zoetmulder MC (1988) *Biochem Soc Trans* **16**: 1081–83.
- 58 Giuseppin MLF, van Eijk HMJ, Bes BCM (1988) *Biotechnol Bioengineering* **32**: 577–83.
- 59 St John TP, Davis RW (1981) *J Mol Biol* **152**: 285–315.
- 60 Vogel K, Hinnen A (1990) *Mol Microbiol* **4**: 2013–18.
- 61 Shuster JR (1989) In *Yeast Genetic Engineering* (Barr PJ, Brake AJ, Valenzuela P, eds) Butterworths, pp 83–108.
- 62 Cousens LS, Shuster JR, Gallegos C, Ku LL, Stempien MM, Urdea MS, Sanchez-Pescador R, Taylor A, Tekamp-Olson P (1987) *Gene* **61**: 265–75.
- 63 Renaud JP, Cullin C, Pompon D, Beaune P, Mansuy D (1990) *Eur J Biochem* **194**: 889–96.
- 64 Janowicz ZA, Melber K, Merckelbach A, Jacobs E, Hartford N, Comberbach M, Hollenberg CP (1991) *Yeast* **7**: 431–43.
- 65 Kaiser CA, Preuss D, Grisafi P, Botstein D (1987) *Science* **235**: 312–17.
- 66 Roggenkamp R, Kustermann-Kuhn B, Hollenberg CP (1981) *Proc Natl Acad Sci USA* **78**: 4466–70.
- 67 Hitzeman RA, Leung DW, Perry LJ, Kohr WJ, Levine HL, Goeddel DV (1983) *Science* **219**: 620–25.
- 68 Wood CR, Boss MA, Kenten JH, Calvert JE, Roberts NA, Emtage JS (1985) *Nature* **314**: 446–49.
- 69 Sato T, Uemura H, Izumoto Y, Nakao J, Nakamura Y, Matsubara K (1989) *Gene* **83**: 355–65.
- 70 Chang CN, Matteucci M, Perry LJ, Wulf JJ, Chen CY, Hitzeman RA (1986) *Mol Cell Biol* **6**: 812–19.
- 71 Singh A, Lugovoy JM, Kohr WJ, Perry LJ (1984) *Nucleic Acids Res* **12**: 8927–38.
- 72 Bitter GA, Chen KK, Banks AR, Lai PH (1984) *Proc Natl Acad Sci USA* **81**: 5330–34.
- 73 Brake AJ, Merryweather JP, Coit DG, Heberlein UA, Masiaz FR, Mullenbach GT, Urdea MS, Valenzuela P, Barr PJ (1984) *Proc Natl Acad Sci USA* **81**: 4642–46.
- 74 Henikoff S, Furlong CE (1983) *Nucleic Acids Res* **11**: 789–800.
- 75 Zaret KS, Sherman F (1984) *J Mol Biol* **177**: 107–35.
- 76 Romanos MA, Scorer CA, Clare JJ (1992) *Yeast* **8**: 423–88.
- 77 Marino MH (1989) *BioPharm* **2**: 18–33.
- 78 Makrides SC (1996) *Microbiol Rev* **60**: 512–38.
- 79 Eckart MR, Bussineau CM (1996) *Curr Opin Biotechnol* **7**: 525–30.
- 80 Meynial-Salles I, Combes D (1996) *J Biotechnol* **46**: 1–14.
- 81 Montreuil J, Vliegthart JFG, Schachter H (1995) *Glycoproteins*, Vol 29a Amsterdam: Elsevier.
- 82 Hagen FK, Van Wuyckhuysse B, Tabak LA (1993) *J Biol Chem* **268**: 18960–65.
- 83 Potvin B, Stanley P (1991) *Cell Regul* **2**: 989–1000.
- 84 LaMont JT, Gammon MG, Isselbacher KJ (1977) *Proc Natl Acad Sci USA* **74**: 1086–90.
- 85 Strous GJ, Berger EG (1982) *J Biol Chem* **257**: 7623–28.
- 86 Kukuruzinska MA, Bergh MLE, Jackson BJ (1987) *Annu Rev Biochem* **56**: 915–44.
- 87 Datar RV, Cartwright T, Rosen CG (1993) *Biotechnology* **11**: 349–57.
- 88 Datta AK, Paulson JC (1995) *J Biol Chem* **270**: 1497–1500.
- 89 Datta AK, Paulson JC (1997) *Indian J Biochem Biophys* **34**: 157–65.
- 90 Mattila P, Joutsjoki V, Kaitera E, Majuri ML, Niittymäki J, Saris N, Maaheimo H, Renkonen O, Renkonen R, Makarow M (1996) *Glycobiology* **6**: 851–59.
- 91 Krezdorn CH, Kleene RB, Watzele M, Ivanov SX, Hokke CH, Kamerling JP, Berger EG (1994) *Eur J Biochem* **220**: 809–17.
- 92 Borsig L, Ivanov SX, Herrmann GF, Kragl U, Wandrey C, Berger EG (1995) *Biochem Biophys Res Commun* **210**: 14–20.
- 93 Malissard M, Berger EG (1998) USGEB98 Lausanne.
- 94 Wen DX, Livingston BD, Medzihradszky KF, Kelm S, Burlingame AL, Paulson JC (1992) *J Biol Chem* **267**: 21011–19.
- 95 Simonen M, Vihinen H, Jamsa E, Arumae U, Kalkkinen N, Makarow M (1996) *Yeast* **12**: 457–66.
- 96 Williams MA, Kitagawa H, Datta AK, Paulson JC, Jamieson JC (1995) *Glycoconjugate J* **12**: 755–61.
- 97 Zimmermann R (1998) *Biol Chem* **379**: 275–82.
- 98 Van Dorst JALM, Tikkanen JM, Krezdorn CH, Streiff MB, Berger EG, Vankuik JA, Kamerling JP, Vliegthart JFG (1996) *Eur J Biochem* **242**: 674–81.
- 99 Berger EG, Malissard M (1997) In *Human Protein Data* (Haeberli A, ed) Weinheim: Wiley-VCH Verlag GmbH.
- 100 Krezdorn CH, Watzele G, Kleene RB, Ivanov SX, Berger EG (1993) *Eur J Biochem* **212**: 113–20.
- 101 Schwientek T, Ernst JF (1994) *Gene* **145**: 299–303.
- 102 Schwientek T, Narimatsu H, Ernst JF (1996) *J Biol Chem* **271**: 3398–405.
- 103 Malissard M, Borsig L, DiMarco S, Grütter MG, Kragl U, Wandrey C, Berger EG (1996) *Eur J Biochem* **239**: 340–48.
- 104 Kleene R, Krezdorn CH, Watzele G, Meyhack B, Herrmann GF, Wandrey C, Berger EG (1994) *Biochem Biophys Res Commun* **201**: 160–67.
- 105 Herrmann GF, Krezdorn CH, Malissard M, Kleene R, Paschold H, Weuster-Botz D, Kragl U, Berger EG, Wandrey C (1995) *Prot Express Purif* **6**: 72–8.
- 106 Borsig L, Berger EG, Malissard M (1997) *Biochem Biophys Res Commun* **240**: 586–89.
- 107 Riederer MA, Hinnen A (1991) *J Bacteriol* **173**: 3539–46.
- 108 Aikawa J, Yamashita T, Nishiyama M, Horinouchi S, Beppu T (1990) *J Biol Chem* **265**: 13955–59.
- 109 Gill GS, Zaworski PG, Marotti KR, Rehberg EF (1990) *Biotechnology* **8**: 956–58.
- 110 Zhu A, Wang ZK, Beavis R (1998) *Arch Biochem Biophys* **352**: 1–8.
- 111 Tsujikawa M, Okabayashi K, Morita M, Tanabe T (1996) *Yeast* **12**: 541–53.
- 112 Watt GM, Lowden PA, Flitsch SL (1997) *Curr Opin Struct Biol* **7**: 52–60.
- 113 Guo Z, Wang PG (1997) *Appl Biochem Biotechnol* **68**: 1–20.
- 114 Elling L (1997) *Adv Biochem Eng Biotechnol* **58**: 89–144.
- 115 Barker R, Olsen KW, Shaper JH, Hill RL (1972) *J Biol Chem* **247**: 7135–47.
- 116 Herrmann GF, Elling L, Krezdorn CH, Kleene R, Berger EG, Wandrey C (1995) *Bioorg Med Chem Lett* **5**: 673–76.

- 117 Berger EG, Greber UF, Mosbach K (1986) *FEBS Lett* **203**: 64–8.
- 118 Unverzagt C (1997) *Carbohydr Res* **305**: 423–31.
- 119 Hokke CH, Zervosen A, Elling L, Joziassse DH, van den Eijnden DH (1996) *Glycoconjugate J* **13**: 687–92.
- 120 Scudder PR, Shailubhai K, Duffin KL, Streeter PR, Jacob GS (1994) *Glycobiology* **4**: 929–32.
- 121 Oehrlein R, Ernst B, Berger EG (1992) *Carbohydr Res* **236**: 335–38.
- 122 Oehrlein R, Hindsgaul O, Palcic MM (1993) *Carbohydr Res* **244**: 149–59.
- 123 Witte K, Sears P, Martin R, Wong CH (1997) *J Am Chem Soc* **119**: 2114–18.
- 124 Ju JM, Kean EL (1992) *Exp Eye Res* **55**: 589–604.
- 125 Schneider R, Hammel M, Berger EG, Ghisalba O, Nuesch J, Gyax D (1990) *Glycoconjugate J* **7**: 589–600.
- 126 Cameron HS, Szczepaniak D, Weston BW (1995) *J Biol Chem* **270**: 20112–122.
- 127 Mas E, Pasqualini E, Caillol N, Elbattari A, Crotte C, Lombardo D, Sadoulet MO (1998) *Glycobiology* **8**: 605–13.
- 128 Murray BW, Takayama S, Schultz J, Wong CH (1996) *Biochemistry* **35**: 11183–95.
- 129 Qiao L, Murray BW, Shimazaki M, Schultz J, Wong CH (1996) *J Am Chem Soc* **118**: 7653–62.
- 130 Romero PA, Lussier M, Sdicu AM, Bussey H, Herscovics A (1997) *Biochem J* **321**: 289–95.
- 131 Roth J (1997) In *The Golgi Apparatus* (Berger EG, Roth J, eds) pp 131–62. Basel: Birkhäuser.
- 132 Chiba Y, Suzuki M, Yoshida S, Yoshida A, Ikenaga H, Takeuchi M, Jigami Y, Ichishima K (1998) *J Biol Chem* **273**: 26298–304.
- 133 Roy SK, Yoko-o T, Ikenaga H, Jigami Y (1998) *J Biol Chem* **273**: 83–90.
- 134 Paulson JC, Colley KJ (1989) *J Biol Chem* **264**: 17615–18.
- 135 Colley KJ, Lee EU, Adler B, Browne JK, Paulson JC (1989) *J Biol Chem* **264**: 17619–22.
- 136 Zigova J, Mahle M, Paschold H, Malissard M, Berger EG, Weuster-Botz D (1999) *Enzyme and Microbiol Technology*, in press.

Accepted 12 March 1999.